Predicting Individual Trip Destinations With Artificial Potential Fields.

Zonta, A.; Smit, S.K.; Haasdijk, Evert

published in
Smart Cities
2017

DOI (link to publisher)
10.1007/978-3-319-59513-9_12

document version
Publisher's PDF, also known as Version of record

Predicting Individual Trip Destinations With Artificial Potential Fields

Alessandro Zonta[1], S.K. Smit[2], and Evert Haasdijk[1]

[1] Vrije Universiteit Amsterdam, The Netherlands
[2] TNO, The Netherlands
{a.zonta,e.haasdijk}@vu.nl,selmar.smit@tno.nl

Abstract. This paper presents a method to model the intended destination of a subject in real time, based on a trace of position information and prior knowledge of possible destinations. In contrast to most work in this field, it does so without the need for prior analysis of habitual travel patterns. The method models the certainty of each POI by means of a virtual charge, resulting in an artificial potential field that reflects the current estimate of the subject's intentions. The virtual charges are updated as new information about the subject's position arrives. We experimentally compare a number of update rules with various parameter settings, showing that it is important to take the distance to a potential destination into account when updating the charge.

Keywords: Human behavior, Intention analysis, Destination Prediction, GPS, Trajectory Database

1 Introduction

Monitoring the movement of individuals offers support for smart decision making in, for example, traffic control, crowd monitoring or location-based services. Insight into the current destination of an individual's movements in particular can provide substantial benefits at a central control level as well as at individual level. At an individual level, location-based services could benefit from such knowledge to reserve resources at their arrival, provide relevant reminders on their to-do lists or suggest places to meet each other [1].

Aggregated knowledge of the destination of individuals in transit can for instance help control rooms to adapt the traffic lights around a city dynamically, or schedule the operation of large infrastructural elements such as ferries or moveable bridges so as to maximise the flow of traffic. In a security or military context, information about the intended destination or itinerary of suspicious individuals or opposing units provides obvious benefits for planning safe transport, (troop) deployment and interception with minimal collateral damage.

With the growing availability of aerial imagery, inexpensive GPS tracking and smartphone usage data, the last few years saw a substantial number of studies that focussed on the use of such data to predict trajectories, waypoints and trip destinations [e.g., 1, 2, 3, 4, 5, 6, 7]. Most work to date analysed data sets of logged trajectories to develop models of habitual movement patterns that can subsequently be employed to predict an individual's movements.

One use of location data is the identification of 'key locations' –destinations or waypoints– that can be used as a basis for predicting an individual's movement. Ashbrook and Starner [1] used clustering approaches to identify key locations from GPS data in an unsupervised manner and subsequently derive a Markov model to predict the probability of transition between these locations. Liao et al. [7] used labelled data to develop a Markov model that extracts and classifies key locations from GPS data. A Bayesian network then models habitual transportation patterns that predict the next location on an individual's path. Scellato et al. [3] identified places that individuals habitually visit based on the length of time a user stayed at a particular position. They then used methods from non-linear time series analysis to predict which of these places an individual is heading for.

A substantial amount of research focusses on habitual traces such as standard paths in an individual's everyday routine. Nicholson and Noble [2] generated a Markov model from smartphone connectivity data that was shown to be capable of accurately predicting individual movement after analysing a week's data. Ziebart et al. [6] analysed data from 25 taxi cab drivers to learn standard routes and developed Markov models to predict the next turn, the itinerary and the destination of a ride. Sadilek and Krumm [5] extracted regularities from each individual's location data and learned their association with particular days of the week. This allowed accurate prediction of the individual's itinerary. Fallis [8] based their predictions of movement on comparisons of an individual's current path with historical traces using a combination of probabilistic inference and path extrapolation.

Lorenzo et al. [9] and Ying et al. [10] used clustering techniques to develop maps that link locations to probable activities. Zheng et al. [11] considered the location data of multiple individuals to identify gatherings and joint movement.

All these approaches have in common that they rely on historical data to identify waypoints and destinations, and to develop models that can eventually predict an individual's destination. Such data is not always available in sufficient quantity, for instance in rural areas, or in one-off situations such as manifestations, festivals or military settings. In this paper, we present a model to predict an individual's destination in real time without the need for prior analysis of habitual travel patterns. The model designates possible destinations ("points of interest" or POIs) as virtual charges that together form an artificial potential field. As location information comes in, the charges are updated so that the POIs that the individual moves towards increase their charge and ones that the individual moves away from decrease. Thus, the charge of the most likely destinations increase, allowing the model to identify an individual's probable target.

Artificial potential fields have been extensively studied as a method in path planning, with obstacles virtually charged so that they repel the subject for obstacle avoidance [e.g., 12, 13, 14] or targets virtually charged so that they attract the subject [e.g., 15, 16].

We experimentally validate and analyse the proposed model in two scenarios. The first uses a simulator that is able to realistically simulate the movement of thousands of people in the city of Rijswijk in the Netherlands [17]. The second scenario uses actual GPS trajectory data collected by Microsoft Research Asia of 182 users over a five-year period [18, 19, 20].

2 Tracking and Prediction system

The basic idea was to build a model able to follow the intention of a tracked subject toward its destination in an urban context. We assume that potential destinations are known in advance; we label them as POIs. Every POI carries a charge q that indicates how likely the POI is to be the subject's target. This defines an Artificial Potential Field (APF) where there are several points, i.e., the POIs, attracting another point, i.e., the tracked person. This APF serves as a model of the subject's destination and intentions and the highest charged POI represents the most likely estimate of the model for the subject's destination.

Equation (1) is used by the model to compute the attraction of the POIs on the tracked person:

$$|E| = \frac{|q|}{d^2} \quad . \tag{1}$$

The formula corresponds to Coulomb's law equation to compute the magnitude of the electric field E created by a charge q at a certain distance d.

To reflect the movement of the subject and use it to adjust the estimate for its destination, an update rule is defined. The update rules are stated in terms of the angle of vision (FOV), the distance from the subject to the POI, the APF and the angle between the POI and the subject's direction of movement. There are two versions of the update rule: (2) uses the angle and (3) also includes the distance.

$$\Delta c = F \cdot (s_1 e^{\beta w_1}) \tag{2}$$

$$\Delta c = F \cdot (s_1 e^{\beta w_1} + \frac{1}{s_2 d^{w_2}}) \tag{3}$$

with $F = 1$ if the POIs are inside the FOV, otherwise $F = -1$. Variables s_1 and s_2 are used to weight the possible influence of β and d to the final result. w_1 regulates how dependent is Δc to the angle β. Small values reduce the influence of the angle and consequently, all the POIs are updated with the same amount even though they have different directions. The variable w_2 is enhancing the difference between the increment of closer and farther POIs. For POIs within the FOV, as for POI$_1$ in Fig. 1, the angle with the edge of the FOV (γ_1 in Fig 1) is used in (2) and (3) to compute Δc. The maximum increase is achieved when a

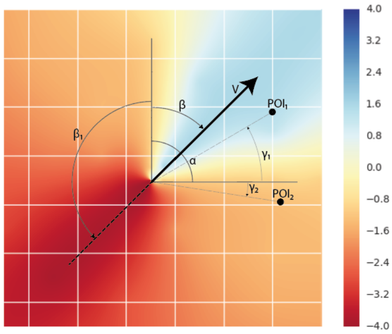


Fig. 1: Representation of virtual charge update. V is the movement direction, α is an algorithm setting and the increment grows following the angles β and β_1. The colours represent the POI's change in charge using (2) with $s_1 = 1$ and $w_1 = 0.01$. γ_1 and γ_2 are used by the model to compute the increment or decrement of the two POIs.

POI is exactly in the direction of the movement. Similarly, for POIs outside the FOV (e.g., POI_2 in Fig. 1) the angle with the edge of the FOV (γ_2 in Fig 1) is used in (2) and (3) to compute the *decrement* that is maximal for POIs directly opposite the movement direction.

Figure 2 shows how the update rule affects the APF to follow the intention of the subject. Figure 2a, 2b, 2c, 2d, and 2e represent the evolution of the APF following the trajectory highlight in black, meanwhile Fig. 2f shows the evolution of the POIs' charge.

Two different methods are considered to compute d and γ. The first option is to use the angle and distance according to the Euclidean Distance (using the Haversine formula) and angle representing the path 'as the crow flies'. The second option is to use a path planner to determine the shortest path towards a POI. In that case, γ is defined as the angle towards the first way-point on the path, and the distance is the length of the planned path towards the POI.

Finally, we have tried a variant where a POI's charge is only updated when the APF is not in line with the current direction of movement.

In the following sections, we use the naming convention N for the rules not using the distance and D for the ones using it and specifically DE if using the Euclidean Distance or DP when using a path planner and finally A indicates that is is only adjusted when not in line with the current forces.

3 Experimental Settings

To test performance and robustness, we have selected two data-sets of trajectories, several different approaches and varied their parameters. The two data-sets differs from each other on, for example, the length of the trajectories, the area

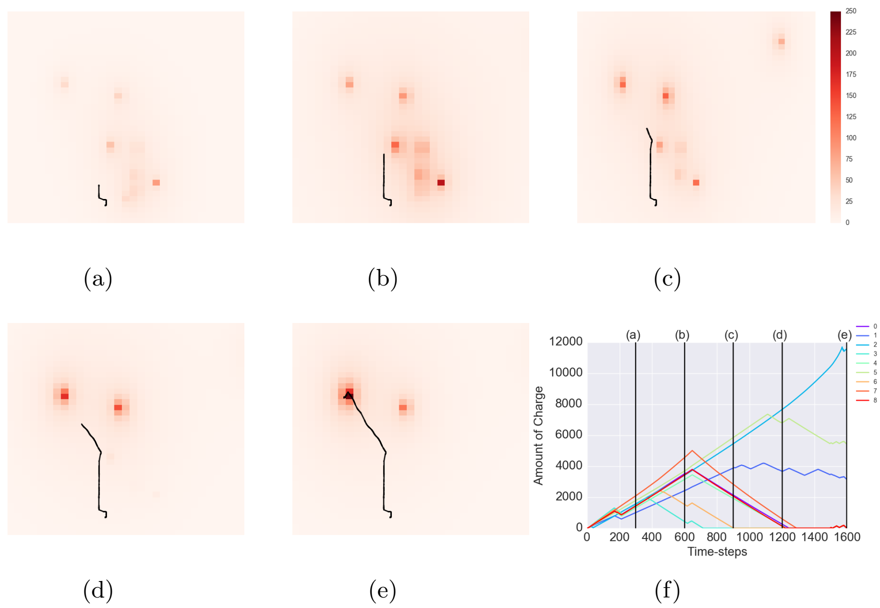
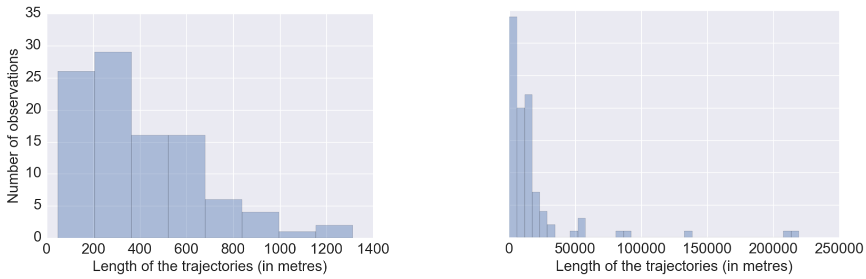


Fig. 2: Representation of how the charges in a Potential Field change. The black line represents a trajectory; the dots describe the intensity of the virtual charge for each POI. Figure 2f shows the time-step where the heat-map are taken, with the charge intensity for each POI.

where the POIs could be found and the transportation systems used. The POIs are determined in the same manner for each dataset: the endpoints of all the trajectories in the dataset are labelled as POIs, and a clustering algorithm is used in order to reduce the possible uncertainty of the locations. The objective was to determine if there is a robust setup that works reasonably well on both of the test-cases using the same parameter values since changing them would not be feasible in an operational setting.

Datasets The first data-set, is created using a simulator of people's movement around a neighbourhood [17]. Each person is generated from open census data and a pattern-of-life, and a social network is created to let behave as normal people with friends and family. The second is a a GPS trajectory data-set collected by Microsoft Research Asia [18, 19, 20] (Geolife Trajectories) tracing 182 users in a period of five years. The data-set contains a broad range of outdoor movement, including life routine as going home or go to work, entertainment and sports activities such as shopping, sightseeing, hiking and cycling. Cars, bus and taxies trajectories are also present in the data-set.

Figure 3 shows how different the two data-sets are in term of trajectories' length.



(a) (b)

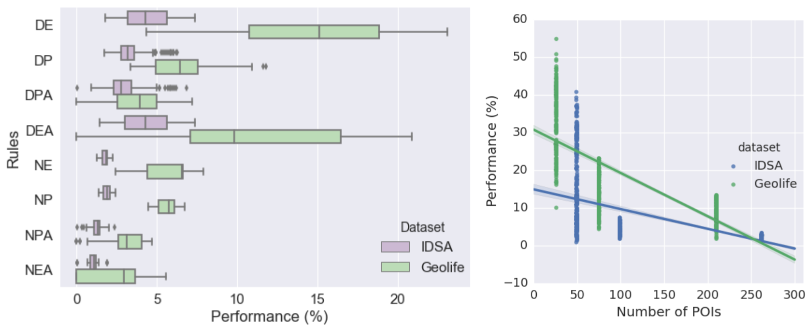
Fig. 3: Length of the two data-sets' trajectories in metres. Fig. 3a represent the IDSA trajectories and Fig. 3b shows the Geolife Trajectories.

Parameters The model has five parameters: the angle of vision and the parameters in (2) and (3). A parameter sweep was performed to find the configuration able to achieve the best prediction. In the following paragraphs, we will use *setting* as a set of parameter. Table 1 shows the parameter tested.

Table 1: Set of parameters tested.

α	w_1	w_2	s_1	s_2
120°	0.005	0.100	0.250	0.100
180°	0.010	0.250	0.500	0.250
240°	0.020	0.500	0.100	0.500

Evaluations The performance is defined as the percentage of the trip length with the target being the highest charged POI in the trajectory. 100 people were tracked with the same setting and the average performance of all of them is considered as the performance of that settings. We use a pairwise non-parametric comparison based on the Wilcoxon-Mann-Whitney test to statistically assess differences in rule performance. We used a Simple Linear Regression Analysis to compute which parameters have most influence on the result. The same trajectories are simulated with three sets of POIs to evaluate how the better rule behaves: firstly, for each subject, a set containing the destinations of all 100 people –the *normal set*– is used. A smaller set is created by removing 50 randomly selected POIs, taking care not to remove the actual target, and a larger set is created by adding 200 further destinations from subjects that are not in the current sample.



(a) (b)

Fig. 4: Figure 4a shows the update rules' performances. The performance is defined as the percentage of the number of time-steps with the target being the highest charged POI in the trajectory. Every bar shows the median and the interquartile range over the median of 100 people tracked per setting. The D on the name correspond to Distance with E for the Euclidean version and P for Path version. A means the usage of the APF and N the angle without the distance. Figure 4b shows the results with different number of POIs with the DE. From the normal set of POIs, some of them are deleted to simulate fewer POIs or different random ones are added to simulate more POIs.

4 Results and Discussion

The median and the interquartile range of the performances per update rule are shown in Fig 4a. Rules based upon the Euclidean distance(DE) achieved the best performance on both data-sets both on median ($p < .001$) and maximum.

On the Geolife data-set, these rules yielded into a performance of over 20% for some of the parameter settings and up to 7% with the IDSA data-set. The inclusion of the 'do-not-always-update'(A) does not improve the results, in fact, it even decreases($p < .001$) the performance. This might be the effect of the not linearity of the trajectories; human behaviour is not regular especially in traffic congested cities. Hence detours from the ideal path a punished more, than moving into the right direction is rewarded.

The decrease in performance when using a path planner was not expected, since in real life people move using pedestrian path and the usage of a path planner seemed a good choice to model those trajectories. However, further analysis shows that the path planner incorporated u-turns rather than going backwards; hence the waypoints of POIs were always in front the tracked person.

It is clear that the number of POIs greatly influences the difficulty of the problem (Figure 4b), but performance remains acceptable.

To investigate the robustness of the different rules regarding their parameter values, response screening is performed (Table 2). As expected, the angle α is

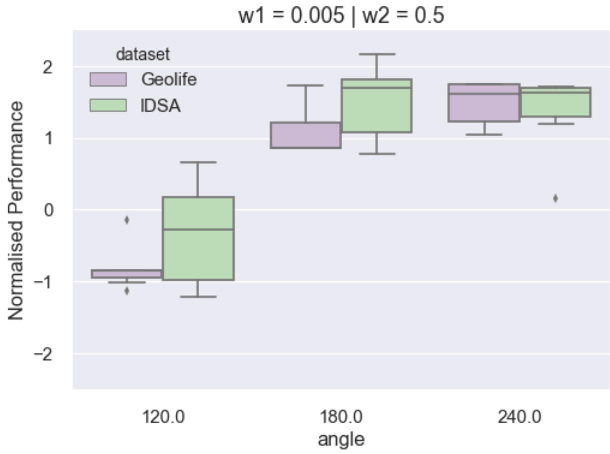


Fig. 5: Performances achieved by *DE* plotted using the best values for the most influencing factors. Every bar shows the median and the interquartile range of the performance over the median of 100 people tracked.

the most influential parameter, but w_2 is also a significant factor. The remaining other variables only have a small impact on performance.

To further investigate robustness, the normalized performances of the best update rule (*DE*), and the most influential parameter (angle) are compared (Figure 5). For both data-sets, the performance increases with the increase of α, indicating that the same parameter-settings can be used for both. The fact that the keep increasing is the effect of how the equations are implemented. α is linearly related with the increment of charge, therefore for big α the Δc is also bigger, and therefore heavily increases the potential, especially if the POI is in the same direction of the movement. This effect also caused the worsened performances for the short trajectories in the IDSA data-set.

Table 2: Impact of the setting's factors to the final results.

Factor	Significance
α	$4.637 \cdot 10^{-26}$
s_1	0.359
s_2	0.220
w_1	0.005
w_2	$2.578 \cdot 10^{-15}$

5 Conclusion and Future Work

We presented a method to model the intended destination of a subject in real time, based on a trace of position information and prior knowledge of possible destinations (points of interest, or POIs). It does so without the need for prior analysis of habitual travel patterns, which puts it in contrast with existing work in this area. The knowledge of POI positions can, however, be derived from historical data, but it may also be available from other sources such as the organisation of one-off events or expert appraisal of the environment (e.g., tactical understanding of terrain in military contexts).

The method models the certainty of each POI by means of a virtual charge, resulting in an artificial potential field that reflects the current estimate of the subject's intentions. The virtual charges are updated as new information about the subject's position arrives. We tested different update rules, finding that it is important to take the distance to a POI into account when updating the charge. Close analysis of well-performing parameter settings for the best update rule considered showed that the best results are achieved when the update rule focusses on increasing the charge of POIs that the subject is moving directly towards.

The method was experimentally validated and analysed on two data-sets, both in an urban environment. One data-set derives from a high fidelity simulation of individuals in an urban setting containing a large concentration of POIs and where trip lengths ranged from hundreds of metres to just over a kilometre. On this data, the model could identify a subject's destination when the trip was over 90% complete on average. The second data-set concerned real-life data, gathered over a five-year period. Trips in this data-set were a lot longer, typically tens of kilometres, ranging up to hundreds of kilometres. Also, POIs were much more sparsely placed. Here, the model could identify the destination more than 20% before the end of the trip. Subsequent analysis bears out that the performance of our method seems to decrease linearly as the number of POIs increases.

While these results are encouraging, the model also has limitations. Mainly, the model struggles with short trips in an environment with many POIs and the accurate identification of a subject's destination is only possible after a substantial part of the trip has passed. Future work will investigate how to improve the update rules, for instance by using modified path planning algorithms to gauge a subject's distance to POIs. Also, prior knowledge of the relative likelihood of POIs and incorporating environmental information can likely improve the performance of the model. The model can further be extended to consider not only the final destination of each trip, but also waypoints on the subject's route. Additional research on different data-sets is needed to ascertain generalisation capability of this approach.

Overall, this paper offers a promising avenue of research towards methods to predict a subject's destination or even itinerary in real time without the need of analysing habitual movement patterns.

Acknowledgement

We thank SURFsara (`www.surfsara.nl`) for the support in using the Lisa Compute Cluster. The research for this paper was financially supported by the Netherlands Organisation for Applied Scientific Research (TNO).

Bibliography

[1] D. Ashbrook and T. Starner. Using GPS to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous Computing*, 7(5):275–286, 2003.

[2] A. J. Nicholson and B. D. Noble. BreadCrumbs: Forecasting Mobile Connectivity. *Proceedings of the 14th ACM international conference on Mobile computing and networking*, (2):46–57, 2008.

[3] S. Scellato, M. Musolesi, C. Mascolo, V. Latora, and A. T. Campbell. NextPlace : A Spatio-Temporal Prediction Framework for Pervasive Systems. *Pervasive Computing: 9th International Conference, Pervasive 2011, San Francisco, USA, June 12-15, 2011. Proceedings*, pages 152–169, 2011.

[4] T. M. T. Do and D. Gatica-Perez. Where and what: Using smartphones to predict next locations and applications in daily life. *Pervasive and Mobile Computing*, 12: 79–91, 2014.

[5] A. Sadilek and J. Krumm. Far Out: Predicting Long-Term Human Mobility. *26th AAAI Conference on Artificial Intelligence*, pages 814–820, 2012.

[6] B. D. Ziebart, A. L. Maas, A. K. Dey, and J. A. Bagnell. Navigate like a cabbie: Probabilistic reasoning from observed context-aware behavior. pages 322–331, 2008.

[7] L. Liao, D. J. Patterson, D. Fox, and H. Kautz. Building Personal Maps from GPS Data. *International Joint Conference on Artificial Intelligence (IJCAI) Workshop on Modeling Others from Observations*, 2005.

[8] A. Fallis. Real-Time Travel Path Prediction Using GPS-Enabled Mobile Phones. *Journal of Chemical Information and Modeling*, 53(9):1689–1699, 2013.

[9] G. D. Lorenzo, S. Phithakkitnukoon, T. Horanont, and G. D. Lorenzo. Activity-Aware Map : Identifying human daily activity pattern using mobile phone data. *Proceedings of the First International Conference on Human Behavior Understanding*, pages 14–25, 2010.

[10] J. J.-C. Ying, W.-C. Lee, T.-C. Weng, and V. S. Tseng. Semantic trajectory mining for location prediction. *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS '11*, page 34, 2011.

[11] K. Zheng, Y. Zheng, N. J. Yuan, S. Shang, and X. Zhou. Online discovery of gathering patterns over trajectories. *IEEE Transactions on Knowledge and Data Engineering*, 26(8):1974–1988, 2014.

[12] Y. K. Hwang and N. Ahuja. A potential field approach to path planning. *IEEE Transactions on Robotics and Automation*, 8(1):23–32, 1992.

[13] A. Howard, M. J. Mataric, and G. S. Sukhatme. Mobile sensor network deployment using potential fields: A distributed, scalable solution to the area coverage problem. *Proceedings of the 6th International Symposium on Distributed Autonomous Robotics Systems (DARS02)*, 5:299–308, 2002.

[14] H. Parunak, L. Purcell, F. Six, N. Station, and M. O'Connell. Digital pheromones for autonomous coordination of swarming uav's. In *1st UAV Conference, Infotech@Aerospace Conferences*, 2002.

[15] R. Mottaghi and R. Vaughan. An integrated particle filter and potential field method applied to cooperative multi-robot target tracking. *Autonomous Robots*, 23(1):19–35, 2007.

[16] H. Helble and S. Cameron. 3-D path planning and target trajectory prediction for the Oxford Aerial Tracking System. *Proceedings - IEEE International Conference on Robotics and Automation*, pages 1042–1048, 2007.

[17] S. de Jong, A. Klein, R. Smelik, and F. van Wermeskerken. Integrating Run-Time Incidents in a Large-Scale Simulated Urban Environment. *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 1401–1402, 2016.

[18] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W.-Y. Ma. Understanding mobility based on GPS data. *Proceedings of the 10th international conference on Ubiquitous computing - UbiComp '08*, (49):312, 2008.

[19] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma. Mining interesting locations and travel sequences from GPS trajectories. *Proceedings of the 18th international conference on World wide web - WWW '09*, (49):791, 2009.

[20] Y. Zheng, X. Xie, and W. Ma. GeoLife: A Collaborative Social Networking Service among User, Location and Trajectory. *IEEE Data Eng. Bulletin*, 33(2):32–40, 2010.